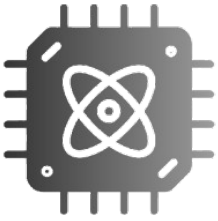
Load Test Report

Pizza Test

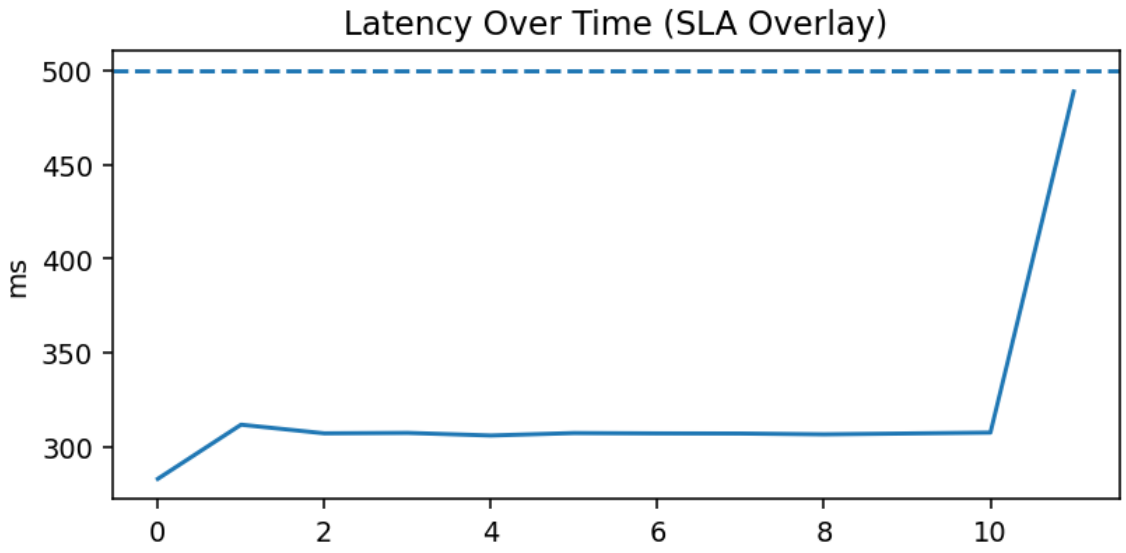
https://quickpizza.grafana.com/

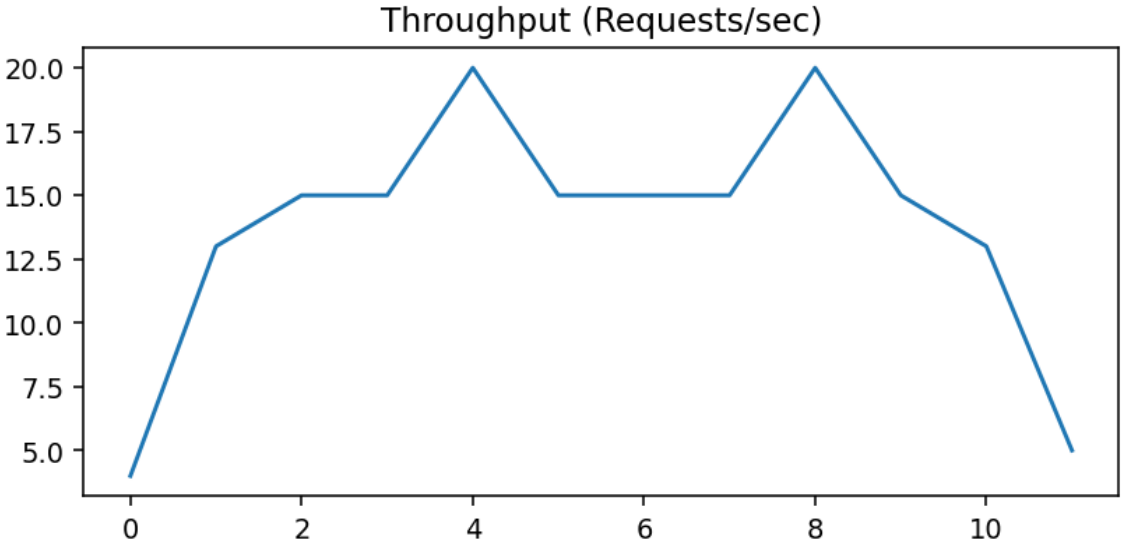
Generated on January 24, 2026 • 18:18

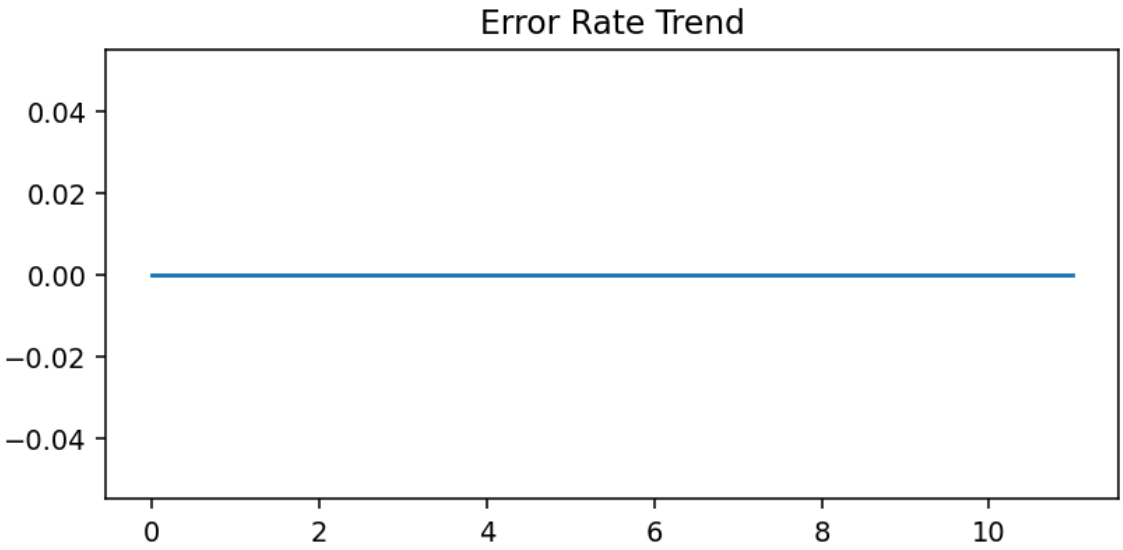


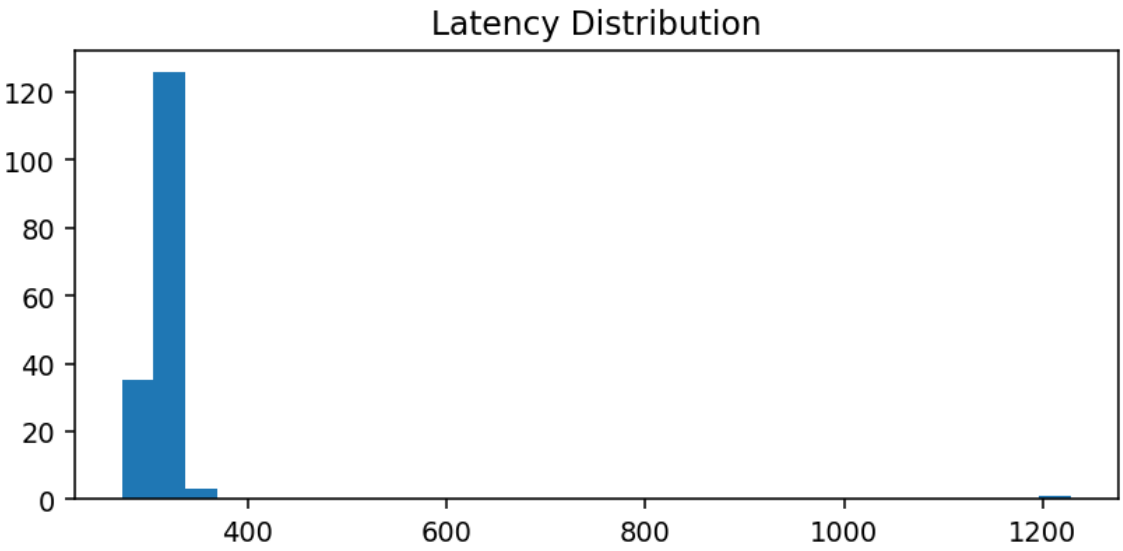
K6 Ai Powered

Performance Dashboard






Error Rate Trend
0.04
0.02
0.00
−0.02
−0.04
0
2
4
6
8
10


Latency Distribution
120
100
80
60
40
20
0
400
600
800
1000
1200

Executive Scorecard

Performance Score	100
SLA Grade	A
Risk Level	Low

Performance Metrics

Avg (ms)	312.21
P95 (ms)	311.39
P99 (ms)	355.43
Requests/sec	15.0
Error Rate	0.0

AI Engineering Analysis

Performance Analysis Report: K6 Load Test Results

Date: October 26, 2023

Prepared For: [Insert Stakeholder Name/Team]

Prepared By: Senior Performance Engineer

---

1. Executive Summary

This report presents an analysis of the provided k6 load test metrics. The overall performance of the tested system appears to be satisfactory under the current load conditions, with a high success rate and no reported errors. The average response time for HTTP requests is within acceptable enterprise thresholds. However, a notable outlier in response time was observed, indicating a potential area for further investigation. The system demonstrated consistent performance across most of the test duration, with minor fluctuations.

---

2. Bottlenecks

Based on the provided data, no significant performance bottlenecks were identified that would broadly impact the system's stability or user experience under the tested load.

However, a single, significant outlier in HTTP request duration (1228.26 ms) was observed at 2026-01-24T18:17:50+00:00. While this is an isolated event and does not represent the typical performance, it warrants attention as it could indicate:

• Transient network issues: A temporary network disruption or congestion.

• Resource contention: A brief spike in resource utilization on a specific server or service.

• Application-level processing delay: A particular request that triggered an unusually long processing time within the application.

• Garbage collection pause: In Java-based applications, a long GC pause could cause such a spike.

---

3. Risk Assessment

The current risk assessment for the tested system, based on these metrics, is Low.

• Error Rate: 0.0% error rate indicates a highly stable and reliable execution of the test.

• Check Success Rate: 100% success rate for all checks confirms that the application is functioning as expected from a functional perspective during the test.

• Response Time Distribution: While the 95th percentile response time is 311.39 ms, which is generally acceptable, the presence of the single outlier (1228.26 ms) introduces a minor, isolated risk. If this outlier were to become more frequent or represent a more common occurrence, it could lead to degraded user experience and potential timeouts for affected users.

---

4. Optimization Recommendations

While the system is performing well, continuous improvement is key. The following recommendations are proposed:

• Investigate the Outlier: Conduct a targeted investigation into the specific request(s) that resulted in the 1228.26 ms response time. This may involve:

• Application Logging: Review application logs for the timestamp of the outlier to identify any errors, warnings, or unusual processing patterns.

• Distributed Tracing: If available, utilize distributed tracing tools to pinpoint the exact service or component that contributed to the extended latency.

• Resource Monitoring: Correlate the outlier event with system resource utilization (CPU, memory, network I/O) on relevant infrastructure.

• Proactive Monitoring: Implement robust monitoring and alerting for response time outliers. Set up alerts for any response times exceeding a predefined threshold (e.g., 2x the 99th percentile) to enable rapid detection and response to future anomalies.

• Code Profiling (if applicable): If the outlier is traced back to a specific application code path, consider profiling that section of code to identify potential inefficiencies.

• Review of http_req_duration max value: The max value of 1228.26 ms is significantly higher than the 99th percentile (355.43 ms). While not a bottleneck in the traditional sense, it indicates that some requests are experiencing considerably longer durations. Understanding the root cause of these infrequent, but long, requests is crucial for ensuring consistent performance.

---

5. Scaling Guidance

The current metrics suggest that the system can handle the tested load effectively. However, scaling guidance should be based on anticipated future demand and the identified areas for improvement.

• Capacity Planning: Based on the current performance, the system appears to have headroom for increased load. However, before scaling, it is crucial to:

• Understand the nature of the outlier: If the outlier is due to a specific, non-scalable operation, scaling might not resolve the underlying issue and could even exacerbate it.

• Define Service Level Objectives (SLOs): Clearly define acceptable response times for different percentiles (e.g., p95, p99) under various load conditions.

• Scaling Strategy:

• Horizontal Scaling: If the system is designed for horizontal scaling, consider increasing the number of application instances or microservices. This is generally the preferred approach for distributed systems.

• Vertical Scaling: If horizontal scaling is not feasible or for specific components, consider increasing the resources (CPU, RAM) of existing instances.

• Load Testing with Increased Load: To validate scaling capabilities, conduct further load tests with progressively higher user loads, focusing on:

• Response Time Degradation: Monitor how response times (especially p95 and p99) change as the load increases.

• Resource Utilization: Observe CPU, memory, and network utilization on all components of the system.

• Error Rates: Ensure that error rates remain at 0% as the load scales.

• Database and External Dependencies: Ensure that any backend databases or external services are also capable of handling the scaled load. Performance bottlenecks in these areas can significantly impact the overall system performance.

This report provides a high-level overview of the performance characteristics observed in the provided k6 metrics. Further detailed analysis and targeted investigations may be required to fully understand and address any potential performance concerns.